# Discrepancies from registered protocols and spin occurred frequently in randomized psychotherapy trials — A meta-epidemiologic study

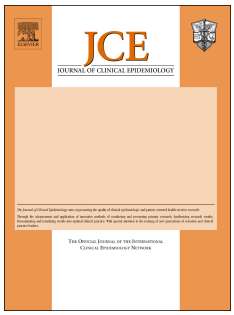
Stoll, Marlene ; Mancini, Alexander ; Hubenschmid, Lara ; Dreimüller, Nadine ; König, Jochem ; Cuijpers, Pim ; Barth, Jürgen ; Lieb, Klaus

Abstract: Objectives: This study aimed to investigate the relationship between trial registration, trial discrepancy from registered protocol, and spin in nonpharmacological trials. Study Design and Setting: Recent psychotherapy trials on depression (2015-2018) were analyzed regarding their registration status and its relationship to discrepancies between registered and published primary outcomes and to spin (discrepancy between the nonsignificant finding in a study and an overly beneficial interpretation of the effect of the treatment). Results: A total of 196 trials were identified, of which 78 (40%) had been registered prospectively and 56 (29%) had been registered retrospectively. In 102 (76%) of 134 registered trials, discrepancies between trial and protocol were present. Of 72 trials with a nonsignificant difference between treatments for the primary outcome, 68 trials (94%) showed spin. Discrepancies from protocol were less frequent in prospectively than in retrospectively registered trials (odds ratio5 0.19; 95% confidence interval [CI]: 0.07-0.52), but regarding the amount of spin, there was no difference between prospectively and retrospectively registered trials (rb=0.12; 95% CI: -0.41 to 0.19) or between registered and unregistered trials (rb=0.22, 95% CI -0.49 to 0.08). Conclusion: Protocol discrepancies and spin have a high prevalence in psychotherapy outcome research. The results show no relation between registration and spin, but prospective registration may prevent discrepancies from protocol.

DOI: https://doi.org/10.1016/j.jclinepi.2020.08.013



Originally published at:
Stoll, Marlene; Mancini, Alexander; Hubenschmid, Lara; Dreimüller, Nadine; König, Jochem; Cuijpers, Pim; Barth, Jürgen; Lieb, Klaus (2020). Discrepancies from registered protocols and spin occurred frequently in randomized psychotherapy trials — A meta-epidemiologic study. Journal of Clinical Epidemiology, 128:49-56.
DOI: https://doi.org/10.1016/j.jclinepi.2020.08.013



# Discrepancies from registered protocols and spin occured frequently in randomized psychotherapy trials – a meta-epidemiologic study

M. Stoll, A. Mancini, L. Hubenschmid, N. Dreimüller, J. König, P. Cuijpers, J. Barth, K. Lieb

PII: S0895-4356(20)30208-0

DOI: https://doi.org/10.1016/j.jclinepi.2020.08.013

Reference: JCE 10254

To appear in: *Journal of Clinical Epidemiology*

Received Date: 4 April 2020

Revised Date: 28 July 2020

Accepted Date: 18 August 2020

Please cite this article as: Stoll M, Mancini A, Hubenschmid L, Dreimüller N, König J, Cuijpers P, Barth J, Lieb K, Discrepancies from registered protocols and spin occured frequently in randomized psychotherapy trials – a meta-epidemiologic study, *Journal of Clinical Epidemiology* (2020), doi: https://doi.org/10.1016/j.jclinepi.2020.08.013.

**Title:** Discrepancies from registered protocols and spin occur frequently in randomized psychotherapy trials – a meta-epidemiologic study

**Authors:** Stoll M[1 2], Mancini A[1], Hubenschmid L[2], Dreimüller N[1], König J[3], Cuijpers P[4], Barth J[5*], Lieb K[1,2*]

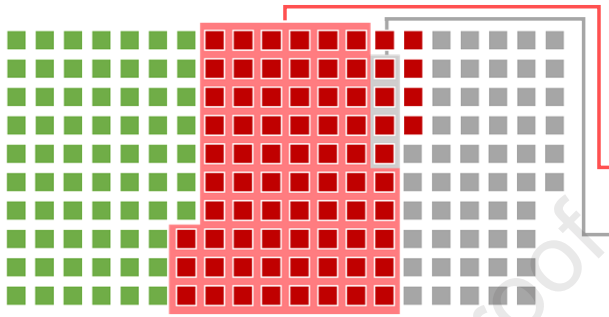
**Affiliations:** [1] Department of Psychiatry and Psychotherapy, University Medical Center Mainz, Untere Zahlbacher Straße 8, 55131 Mainz, Germany
[2] Leibniz Institute for Resilience Research, Wallstraße 7, 55122 Mainz, Germany
[3] Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Mainz, Obere Zahlbacher Straße 69, 55131 Mainz, Germany
[4] Department of Clinical, Neuro, and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Van der Boechorstraat 7, 1081 BT Amsterdam, the Netherlands
[5] Institute for Complementary and Integrative Medicine, University Hospital Zurich and University Zurich, Sonneggstraße 6, 8091 Zurich, Switzerland
*Both authors contributed equally.

**CRediT author statement:**

**Stoll Marlene**: Conceptualization, Methodology, Investigation, Writing Original Draft, Writing – Review & Editing, visualization, Project administration
**Mancini Alexander**: Methodology, Investigation, Writing – Review & Editing, Formal Analysis
**Hubenschmid Lara**: Investigation, Writing – Review & Editing, Project Administration
**Dreimüller Nadine**: Writing – Review & Editing
**König Jochem**: Formal analysis, Writing – Review & Editing
**Pim Cuijpers**: Conceptualization, Resources, Data Curation, Writing – Review & Editing
**Barth Jürgen**: Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision
**Lieb Klaus**: Conceptualization, Methodology, Writing – Original Draft, Writing – Review & Editing, Supervision, Funding Acquisition

196 psychotherapy trials were included.

**Assessment of trial effectiveness:**
The investigated treatment is …
**effective (67/196, 34%)**
**not effective (77/196, 39%)**
**unclear (52/196, 27%)**
… on depression.



**Assessment of spin:**
In trials with non-significant outcomes, results were reported with …
**at least one form of spin (68/72, 94%)**
**no spin at all (4/72, 6%)**
*5 trials were not rateable*

Prevalence of **registration**:
134/196 trials (68%) had registered protocols; 62/196 trials (32%) had no protocol.

Median amount of spin per trial:
5.75

→ Median amount of spin in registered vs. unregistered trials:
5 vs. 7
$r_b$ = -.22
95% CI [-.49, .08]

→ Median amount of spin in prospectively vs. retrospectively registered trials:
5 vs. 6
$r_b$ = -.12
95% CI [-.41, .19]

**Assessment of protocol discrepancies:**
Trial is …
**discrepant (102/134, 76%)**
**concordant (32/134, 24%)**
… with the protocol.

→ Discrepancies in retrospectively vs. prospectively registered trials:
51/56 (91%) vs. 51/78 (65%)
*OR* = 0.19
95% CI [0.07, 0.52]

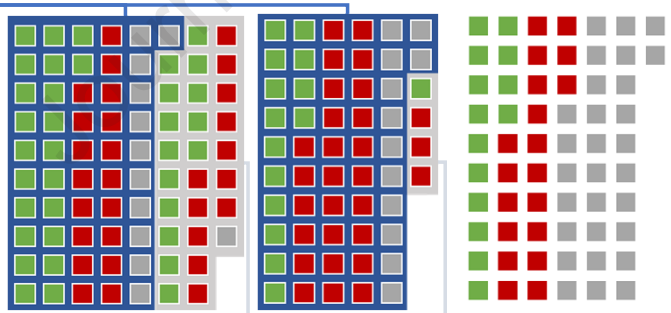
Time point of registration:

**Prospective**
78/134 (58%)

**Retrospective**
56/134 (42%)

No registration:

*Discrepancies not traceable*



**Relation**:
registration and
- protocol discrepancies
- spin

- **Retrospectively** registered trials are **more likely** to show **discrepancies** than prospectively registered trials.
- **No difference** was found between registered and unregistered trials regarding **spin**.
  **No difference** was found between prospectively and retrospectively registered trials regarding **spin**.

**Title**

# Discrepancies from registered protocols and spin occured frequently in randomized psychotherapy trials – a meta-epidemiologic study

**Authors**

Stoll M[1,2], Mancini A[1], Hubenschmid L[2], Dreimüller N[1], König J[3], Cuijpers P[4], Barth J[5*], Lieb K[1,2*]

**Affiliations**

[1] Department of Psychiatry and Psychotherapy, University Medical Center Mainz, Untere Zahlbacher Straße 8, 55131 Mainz, Germany

[2] Leibniz Institute for Resilience Research, Wallstraße 7, 55122 Mainz, Germany

[3] Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Mainz, Obere Zahlbacher Straße 69, 55131 Mainz, Germany

[4] Department of Clinical, Neuro, and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Van der Boechorstraat 7, 1081 BT Amsterdam, the Netherlands

[5] Institute for Complementary and Integrative Medicine, University Hospital Zurich and University Zurich, Sonneggstraße 6, 8091 Zurich, Switzerland

*Both authors contributed equally.

Corresponding author

Jürgen Barth, Institute for Complementary and Integrative Medicine, University Hospital Zurich and University Zurich, Sonneggstraße 6, 8091 Zurich, Switzerland.

**Word count: 3.558 / 3.000**

**Word count Abstract: 200 / 200**

**References: 42**

**ABSTRACT**

**Objective:** To investigate the relationship between trial registration, trial discrepancy from registered protocol and spin in non-pharmacological trials.

**Study Design and Setting:** Recent psychotherapy trials on depression (2015 – 2018) were analyzed regarding their registration status and its relationship to discrepancies between registered and published primary outcomes and to spin (discrepancy between the non-significant finding in a study and an overly beneficial interpretation of the effect of the treatment).

**Results:** 196 trials were identified of which 78 (40%) had been registered prospectively and 56 (29%) retrospectively. In 102 (76%) of 134 registered trials, discrepancies between trial and protocol were present. Of 72 trials with a non-significant difference between treatments for the primary outcome, 68 trials (94%) showed spin. Discrepancies from protocol were less frequent in prospectively than in retrospectively registered trials (*OR* = 0.19; 95% CI [0.07, 0.52]), but regarding the amount of spin there was no difference between prospectively and retrospectively registered trials ($r_b$ = -.12; 95% CI [-.41;.19]) or between registered and unregistered trials ($r_b$ = -.22; 95% CI [-.49;.08]).

**Conclusion:** Protocol discrepancies and spin have a high prevalence in psychotherapy outcome research. Results show no relation between registration and spin, but prospective registration may prevent discrepancies from protocol.

**Keywords:** psychotherapy, depression, reporting bias, spin in research, conflict of interest, review

**Running title:** Protocol discrepancies / spin in psychotherapy studies

**Word count: 200**

**What is new?**

- *Key Findings*

Discrepancies from the registered protocol manifest in 76% of psychotherapy trials and nearly all trials with non-significant effects between treatments show some form of spin

- *What does this add to what was known?*

Protocol discrepancies are less frequent in psychotherapy trials which are registered prospectively as compared to retrospectively registered trials

- *What is the implication and what should change now?*

Spin is not prevented by registration of trials.

Policies such as reporting guidelines should be promoted by relevant stakeholders.

## 1 Introduction

Since 2013, the World Medical Association declaration of Helsinki states that “every research study involving human subjects must be registered in a publicly accessible database before recruitment of the first subject” [1]. In medicine as well as in other disciplines such as psychology, the (pre-)registration of clinical trials or studies in public registries is seen as an effective tool to improve conduct, analysis and interpretation of studies and to increase reliability and consequently credibility of research [2–5]. To publish a study protocol in public registries includes the registration of study characteristics such as study design, main outcomes and main analyses [6]. Registration of studies thereby is intended to prevent selective publication and selective reporting of outcomes [3, 7], to prevent unnecessary double research effort, and to give patients, the public and other stakeholders such as ethical review boards an overview of ongoing or planned trials [7]. In the medical sciences, up to 90% of clinical trials are registered, but only 60–77% are registered prospectively, i.e. before enrolling the first participant [8, 9]. These numbers are based mainly on pharmacological trials published in journals that endorse the International Committee of Medical Journal Editors’ (ICMJE) guidelines [7]. Registration of non-pharmacological trials has not been well investigated yet. One study of psychotherapy trials found that 60% of the investigated trials had been registered, 24% of them prospectively [10]. More evidence is available for health intervention trials which include pharmacological as well as non-pharmacological trials with about half of them being registered [11–15].

If there are systematic discrepancies between the information given in published trials and their respective registered protocols, this is called outcome reporting bias. Outcome reporting bias is frequently found (28–62%) in pharmacological as well as non-pharmacological trials [14, 16-19]. Another form of bias in the reporting of results can occur if the publications themselves show an overly beneficial interpretation of the reported effect of the treatment [20]. This interpretation bias is called “spin” and is also highly prevalent in the biomedical literature [20, 21]. It has been estimated that 56% of the abstracts published in psychiatry

and psychology journals contain spin [22]. Different biases can accumulate and interact: Trials with positive outcomes are more likely to be published (publication bias) and significant outcomes are more likely to be included in a published trial, while negative outcomes are changed or omitted (another form of outcome reporting bias). In case negative outcomes are reported, they are sometimes reported in an overly beneficial way (spin). These biases may arise in conjunction and thereby leave non-significant results out of eyeshot, which creates a risk of a distorted image of the actual evidence in the published literature [23].

The field of psychotherapy research is of particular interest regarding reporting biases: Contrary to most medical trials, a typical psychotherapy trial is conducted by a researcher with clinical expertise who works as a therapist and whose school of thought is exceptionally shaped by a long education in this therapy [24]. While in medical research the pharmaceutical industry as an external factor may play a relevant role in the conduction of trials, industry is less involved in the conduction of psychotherapy trials. Therefore, psychotherapy trials are more dependent on the individual researcher and the researcher's personal interests in the outcome of the trial might play a more important role. These interests are discussed in terms of researcher allegiance. Evidence shows that researchers with higher researcher allegiance often published studies with larger effects [25].

To the best of our knowledge, no studies have investigated the registration of psychotherapy randomized trials and its relationship to both protocol discrepancies and spin. We, therefore, investigated in the present study registration status and registration time point, discrepancies between trials and protocols, and spin in a larger number of psychotherapy trials. The objectives of the present study were (1) to investigate the extent to which recent psychotherapy trials on depression are prospectively or retrospectively registered; (2) to investigate the respective prevalence of protocol discrepancies and spin; and (3) to examine the relationship of registration status and registration time point to protocol discrepancies and spin. Further, the relationship between protocol discrepancies and trial effectiveness was

explored to provide preliminary evidence for the prevalence of outcome reporting bias in trials with non-pharmacological interventions.

## 2 Methods

### Selection of trials

Trials were retrieved from a collection of psychotherapy trials on depression provided by Cuijpers et al. [26]. We focused on one particular disorder to minimize variation and decided upon depression because it is a highly prevalent disorder [27] with a high level of disease burden [28], leading to an ongoing development of treatments, and therefore a large evidence base for research on treatments is available [29]. In the trial collection by Cuijpers et al. [26], studies were eligible to be included if they investigated the treatment of a depressive disorder or an elevated level of depressive symptomatology and at least one treatment arm is psychological and for adults. The systematic literature search is updated every year and was conducted in the databases PubMed, Embase, PsycINFO and Cochrane Register of Controlled Trials up to January 1, 2019, with no language restrictions. For more details regarding the database, see Cuijpers (2017) [29], Cuijpers et al. (2008) [30] and Cujipers (2016) [31]. Specific inclusion criterion for the present study was that randomized trials of the described collection were published between January 1, 2015 and December 31, 2018. Studies published before 2015 were excluded to get a current picture that is not interfered by earlier standards.

### Data extraction

For each trial, two independent reviewers (AM, MS) extracted information and conducted assessments for protocol discrepancies and spin. Disagreements were resolved by discussion and consultation with a third investigator (JB), if necessary. Reviewers extracted the following items from the trial protocol and/or the published manuscript: registration

number and time point of registration, primary outcomes, and statistical significance on the published primary outcome. If available, the definition of the primary outcome was extracted, including measurement scale, time point and time frame (i.e. involving "baseline"-time point). To find the respective information, texts were searched manually.

**Assessing trial registration status**

To assess trial registration status, the trial was screened to identify a registration number. If none was found, we searched online registries in the following order by using the surname of the first author, the name of the treatment and mental health condition as search parameters: ClinicalTrials.gov, www.who.int/trialsearch, www.isrctn.com, and author's local registry. If a registration number and the respective registration were found, the trial was rated as *registered*. Trial registration search was conducted in October 2019.

**Assessing data in the registrations**

Trials were considered *prospectively registered* if the registration date preceded participant enrollment date or if the trial was registered within one month of participant enrollment (e.g. participant enrollment date: May 1; registration date: May 15). Trials were considered *retrospectively registered* if the trial was registered more than one month after participant enrollment had begun or if registration within one month was unclear (e.g. participant enrollment date: May; registration date: June) or if participant enrollment began in the same year the trial was registered and no specific month was specified.

Registrations were screened for the definition (method of measurement, time point or frame) of the primary outcome. If the primary outcome definition had been changed, we extracted information of the "original" rather than a changed version of the primary outcome definition. The registered primary outcome was deemed to be defined *exactly* if one scale and one time point was defined and *inexactly* if none or more than one scale or none or more than one time point or a time frame was defined.

**Assessing data in the publications**

We screened publications for the terms “primary outcome(s)”, “primary endpoint(s)”, “main outcome(s)”, or “main endpoint(s)”. If mentioned, the respective definition (i.e. method, time point, or time frame of measurement) was extracted. For this purpose, we screened the paragraph where the term was mentioned and extracted the information regarding measurement and time point that was semantically the closest. If more than one primary outcome was defined, we extracted all mentioned primary outcomes.

Next, we assessed the exactness of the definition for every mentioned primary outcome (see above) and its statistical significance. A primary outcome was regarded as *significant* if it was exactly defined and the statistical test reached $p < .05$ (if not otherwise specified), and as *non-significant* if this test did not reach statistical significance. In case of an inexactly defined primary outcome, we extracted the statistical significance of this primary outcome at posttest in favor of the intervention group.

Afterwards, we assessed effectiveness of trials. A trial was classified as *effective*, if the reported primary outcome was statistically significant or, in case there was more than one primary outcomes per trial, all of them were statistically significant. It was classified as *not effective*, if at least one primary outcome was statistically non-significant.

**Analysis of protocol discrepancies**

We analyzed protocol discrepancies in all trials for which a registration could be identified by firstly, comparing the respective registered and published primary outcomes. They were classified as *discrepant* if their definitions differed (e.g., different methods of measurement) or if the amount of information differed (e.g., a time point was registered but not reported). They were classified as *concordant* if the registered and reported primary outcomes matched exactly. This was also assumed if the time point of the reported primary outcome was within the time frame of the registered primary outcome.

In a second step, we assessed protocol discrepancy per trial which was considered present if there was a discrepancy between the registered and the reported primary outcome; and if

there was more than one reported primary outcome, protocol discrepancy per trial was considered present if there was at least one discrepancy between registered and reported primary outcomes in the trial. Protocol discrepancy per trial was considered non-present if the primary outcome or, in case of more than one primary outcome, all of them were reported exactly as they had been registered. Notably, this approach to assess protocol discrepancies differs from assessments of outcome reporting bias [17].

**Assessing spin**

To assess spin, we examined all trials with at least one non-significant PO. We adapted the coding manual by Gewandter et al. (2015) [32] which they had developed based on Boutron et al. (2010) [20]. Seven forms of spin were investigated: 1. selective reporting (the non-significant primary outcome is not mentioned in the screened section), 2. distracting with secondary analyses (the primary outcome is not mentioned but significant secondary analyses are), 3. distracting with within-group differences (the primary outcome is not mentioned but significant within-group differences are), 4. focus on significant secondary analyses (a. secondary analyses are mentioned before the primary outcome; b. effect sizes are mentioned instead of primary outcome effect sizes; c. effect is depicted in figures but primary outcome is not), 5. focus on significant within-group differences over time (a.-c., see above), 6. interpreting non-significant primary results as showing treatment equivalence in a superiority trial, 7. claiming or emphasizing the beneficial effect of the treatment despite a non-significant outcome. Spin forms were investigated in five sections of the publication: abstract results and conclusions, main text results, discussions, and conclusions (see table 2). Additionally to the assessment of the single spin forms in the different publication sections, we calculated the amount of spin per trial by summing up the occurrence of spin forms in that trial. To estimate interrater reliability of the spin scores per trial, we calculated the intraclass correlation coefficient (ICC), which was $ICC = 0.81$, based on a mean-rating ($k = 2$), absolute-agreement, 2-way random effects model. If two or more primary outcomes

per trial were non-significant, we assessed spin separately per primary outcome, and the amount of spin per trial was generated by using the average score.

**Statistical analyses**

In addition to the descriptive analysis of the sample characteristics, we aimed to investigate the relationship between registration time point and protocol discrepancy; registration status and amount of spin; and registration time point and amount of spin. Further, we investigated the relationship between registration status and effectiveness; registration time point and effectiveness; and protocol discrepancy and effectiveness. For quantification of the relationship between two binary characteristics, we calculated Odds Ratios (OR) and corresponding 95% confidence intervals (CI). For quantification of the relationship between a binary characteristic and the amount of spin, we calculated Mann-Whitney-U-tests and provide rank biserial correlations ($r_b$) with 95% CI. Statistics of the distribution of the amount of spin are reported as median and quartiles (interquartile range: IQR).

## 3 Results

**Selection of eligible RCTs**

204 trials of the database matched the inclusion criteria. Eight trials had to be excluded because they were duplicates in the database. Finally, we extracted and analyzed 196 trials (characteristics of each trial are shown in eTable1).

**Registration status, registered and reported primary outcomes**

134 (68%) of 196 trials had been registered in a clinical trial registry. Of those, 78 (58%) had been registered prospectively and 56 (42%) retrospectively (table 1A). In the 134 protocols, 197 primary outcomes were registered which are on average 1.47 registered POs per protocol. Only 26% of them were exactly defined.

- Insert Table 1 about here -

In the 196 published trials, 194 primary outcomes were reported of which 89 (46%) were statistically significant (table 1B). In 46 (23%) of 196 published trials, no primary outcome was defined. Overall, we could assess effectiveness in 144 of the 196 trials, and classified 67 (47%) as effective and 77 (53%) as not effective.

**Prevalence of protocol discrepancies**

At least one discrepancy between protocol and trial was present in 102 (76%) of 134 registered trials. Prevalence of protocol discrepancies were lower in prospectively (51/78, 65%) than in retrospectively (51/56, 91%) registered trials. Odds of protocol discrepancies were significantly reduced with prospective registration (*OR* = 0.19; 95% CI [0.07, 0.52]).

Registered and reported primary outcomes and rating of protocol discrepancies in all 196 included trials are shown in eTable 2.

**Prevalence of spin**

72 trials had at least one non-significant primary outcome and were assessed for spin. Spin forms and amount of spin for each of those 72 trials are listed in eTable 3a and 3b. 68 trials (94%) showed at least one form of spin (median amount of spin per trial was 5.75, *IQR* 3–8). As shown in table 2, the most frequently used forms of spin were that the non-significant PO was not mentioned in the abstract conclusion section (selective reporting in 37/69 trials; 54%), and that the beneficial effect of the treatment was claimed in the abstract (30/69 trials; 43%) or the main text (22/48 trials; 46%) conclusion section. The text section with the highest

prevalence of spin ratings was the abstract conclusion section, in which 56 of 69 (81%) investigated trials showed some form of spin, and the main text discussion section, in which 58 of 72 (81%) investigated trials showed some form of spin.

- insert Table 2 about here -

We could assess spin in 28 prospectively registered trials, 25 retrospectively registered trials and 19 unregistered trials. The percentages of trials that showed spin were 89%, 96% and 100%, respectively. Amount of spin per trial did not differ significantly between unregistered (*Mdn* = 7; *IQR* 4–8) and registered trials (*Mdn* = 5; *IQR* 3–8; $r_b$ = -.22; 95% CI [-.49,.08]) nor between retrospectively (*Mdn* = 6; *IQR* 3–8) and prospectively registered trials (*Mdn* = 5; *IQR* 2.63–8; $r_b$ = -.12; 95% CI [-.41,.19]).

In 4% of investigated trials (3/72), spin was found in only one section of the trial publication, in 10% (7/72) in two, in 28% (20/72) in three, in 31% (22/72) in four, and in 22% (16/72) in five sections of the trial publication.



**Effectiveness of trials**

Trials reported effective interventions in 67 (47%) of 144 trials that unambiguously reported about effectiveness. The prevalence rates of effective trials are: registered vs. unregistered trials: (53/111, 48% vs. 14/33, 42%), prospectively vs. retrospectively registered trials (36/66, 55% vs. 17/45, 38%), and trials with vs. trials without protocol discrepancies (36/80, 45% vs. 17/31, 55%).

The odds to report about an effective treatment did not significantly increase with registered compared to unregistered trials (*OR* = 1.24; 95% CI [0.57, 2.72]) or with prospective compared to retrospective registration (*OR* = 1.98; 95% CI [0.91, 4.28]). The odds to report

about an effective treatment did not significantly decrease with trials that showed protocol discrepancies compared to trials without protocol discrepancies (*OR* = 0.67; 95% CI [0.29, 1.55]).

## 4 Discussion

### Principal findings

Of 196 trials, 40% had been registered prospectively and 29% retrospectively. Protocol discrepancies were present in 75% of registered trials. Non-significant primary outcomes were interpreted with some form of spin in 94% of all trials. We found no differences in protocol discrepancies and spin between registered and unregistered trials, but protocol discrepancies were less likely if trials were registered prospectively.

### Findings in context

Our finding of 68% registered trials is in the range of previous research on clinical trial registration of health interventions [10, 14, 15], but prevalence of registration appears lower than in comparable studies on pharmacological trials where clinical trial registration is an established standard [8, 9]. To compare our findings regarding outcome reporting bias with other studies is difficult because the definition of outcome reporting bias often differs between studies. Our approach was different, since we measured discrepancies between registered and published primary outcomes in a first step and then related protocol discrepancy to trial effectiveness in a second step, whereas other studies assessed outcome reporting bias as discrepancy that favors the published primary outcomes [14, 16, 17]. Our findings support the conclusion of other researchers that registrations of pharmacological and non-pharmacological trials are of low quality (e.g., high prevalence of inexactly defined primary outcomes) and that registered and published information often is discordant [10, 11, 18, 19]. Regarding spin, our results are in line with previous research in that most spin was found in the conclusion sections of the investigated trials [20, 22, 33]. More detailed

comparisons are difficult due to different definitions of spin. Further studies should find a standardized spin measurement method and make direct comparisons, e.g. between pharmacological and non-pharmacological trials. The reasons for the high prevalence of reporting biases such as spin are still unclear. Chiu et al. (2017) [21] showed that funding source is one of the most frequently investigated factors associated to spin, but they did not find a significant association between industry sponsorship and spin. It might be speculated, especially for psychotherapy trials, that other factors such as researcher allegiance [24, 25, 34, 35] or inappropriate incentives fueled through the academia reward system may contribute to the high prevalence of bias [36, 37].

**Strengths and limitations**

A strength of the present study is that registration status, protocol discrepancies, spin and their interactions were investigated for the first time in a large sample of non-pharmacological outcome studies. A further strength was that we investigated spin in a very detailed way, with an extension of the scale developed by Gewandter et al. (2015) [32]. Limitations are: First, protocol discrepancies obviously could only be investigated in registered trials. We therefore do not know whether registration per se is effective to reduce the risk of outcome reporting bias. Second, our analysis of protocol discrepancies and spin was in general only possible if outcomes were (adequately) reported, which often was not the case.

**Conclusions and policy implications**

The high prevalence of protocol discrepancies and spin and the fact that only the risk of protocol discrepancies, but not of spin, was reduced by prospective registration, suggest that more effective ways than mere registration of studies are needed to increase the trustworthiness of psychotherapy outcome research.

First, better adherence to standards is needed. Many journals that publish non-pharmacological trials do not have policies that require prospective registration. The rate of prospectively registered trials published in these journals is much lower than in journals that endorse such guidelines [12, 38]. The most well-known reporting guideline is the Consolidated Standards of Reporting Trials 2010 Statement (CONSORT) that requires that primary and secondary outcome measures are completely defined, including “how and when they were assessed”, that trial`s registration number and name of trial registry are reported and that any changes of outcomes after the trial commenced are mentioned, “with reasons” [39]. We especially encourage journals publishing non-pharmacological trials to implement these guidelines. Reporting guidelines should also be promoted by graduate schools and writing courses and they should receive appropriate attention by scientific boards of academic associations. Second, other potential sources of bias despite funding, e.g. researcher allegiance [34, 36], have to be better identified and need transparency. Third, the publication format of registered reports, which is adopted by an increasing number of journals [37, 40], should be promoted. In such registered reports, study protocols are submitted to a journal before any data is gathered or analyzed. Study protocols then undergo a peer-review process and, with acceptance, the publication of the trials’ results after data collection, analysis and interpretation is guaranteed independent from the finding [41, 42]. To enhance acceptance among researchers, however, a fast reviewing process will be of high importance.

In conclusion, this study shows a high prevalence of protocol discrepancies and spin and low rates of registered trials in psychotherapy research. Prospective registration in this sample was associated with less protocol discrepancies than retrospective registration.

**Funding**

This work was funded by the Volkswagen Foundation [grant no. A118085, ref. 91498 to KL]. The Foundation did not take role in the study design, data collection, analysis or interpretation, or in writing of the report of the submitted article. All researchers are independent from funders and all authors had full access to the data.

**Declarations of interest**

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf. MS, AM and LH declared that they had received salary from the Volkswagen Foundation. KL declared that he received a research grant by Volkswagen Foundation. However, the funder had no influence on the study findings and interpretation. JB, PC, ND and JK have nothing to disclose. All authors declared: no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced submitted work.

**Data statement:** all data are available in the online appendix (eTables 1–3)



**Acknowledgements**

We thank Marc Himmelmann, Janosch Weisskircher and Cora Koch for their help and input in the early stages of the project. We also thank Ioana Cristea and Whitney Yoder for their feedback on our research project and support regarding the database.

**References**

[1] World Medical Association. WMA declaration of Helsinki – ethical principles for medical research involving human subjects. [Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ accessed June 2020]

[2] Allen C, Mehler DMA. Open science challenges, benefits and tips in early career and beyond. PLOS Biology 2019;17(5):e3000246. doi: 10.1371/journal.pbio.3000246

[3] Dal-Ré R, Bobes J, Cuijpers P. Why prudence is needed when interpreting articles reporting clinical trial results in mental health. Trials 2017;18(1):143-43. doi: 10.1186/s13063-017-1899-2

[4] DeAngelis CD, Drazen JM, Frizelle FA, et al. Clinical trial registration: A statement from the International Committee of Medical Journal Editors. Archives of Dermatology 2005;141(1):76-77. doi: 10.1001/archderm.141.1.76

[5] Ioannidis JPA. Why most published research findings are false. PLOS Medicine 2005;2(8):e124. doi: 10.1371/journal.pmed.0020124

[6] World Health O. International standards for clinical trial registries: the registration of all interventional trials is a scientific, ethical and moral responsibility. version 3.0 ed. Geneva: World Health Organization 2018:48

[7] ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals [Available from: http://www.icmje.org/icmje-recommendations.pdf accessed March 2020]

[8] Gopal AD, Wallach JD, Aminawung JA, et al. Adherence to the International Committee of Medical Journal Editors' (ICMJE) prospective registration policy and implications for outcome integrity: a cross-sectional analysis of trials published in high-impact specialty society journals. Trials 2018;19(1):448. doi: 10.1186/s13063-018-2825-y

[9] Huser V, Cimino JJ. Evaluating adherence to the International Committee of Medical Journal Editors' policy of mandatory, timely clinical trial registration. Journal of the American Medical Informatics Association 2013;20(e1):e169-74. doi: 10.1136/amiajnl-2012-001501 [published online first: 2013/02/12]

[10] Bradley HA, Rucklidge JJ, Mulder RT. A systematic review of trial registration and selective outcome reporting in psychotherapy randomized controlled trials. Acta Psychiatrica Scandinavica 2017;135(1):65-77. doi: 10.1111/acps.12647 [published online first: 2016/10/01]

[11] Azar M, Riehm KE, McKay D, et al. Transparency of outcome reporting and trial registration of randomized controlled trials published in the journal of consulting and clinical psychology. PLOS ONE 2015;10(11):e0142894-e94. doi: 10.1371/journal.pone.0142894

[12] Azar M, Riehm KE, Saadat N, et al. Evaluation of journal registration policies and prospective registration of randomized clinical trials of nonregulated health care interventions. JAMA Internal Medicine 2019;179(5):624-32. doi: 10.1001/jamainternmed.2018.8009

[13] Cybulski L, Mayo-Wilson E, Grant S. Improving transparency and reproducibility through registration: The status of intervention trials published in clinical psychology journals. Journal of Consulting and Clinical Psychology 2016;84(9):753-67. doi: 10.1037/ccp0000115 [published online first: 2016/06/10]

[14] Scott A, Rucklidge JJ, Mulder RT. Is mandatory prospective trial registration working to prevent publication of unregistered trials and selective outcome reporting? An observational study of five psychiatry journals that mandate prospective clinical trial registration. PLOS ONE 2015;10(8):e0133718. doi: 10.1371/journal.pone.0133718

[15] Shinohara K, Tajika A, Imai H, et al. Protocol registration and selective outcome reporting in recent psychiatry trials: new antidepressants and cognitive behavioural therapies. Acta Psychiatrica Scandinavica 2015;132(6):489-98. doi: 10.1111/acps.12502

[16] Dwan K, Altman DG, Cresswell L, et al. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database of Systematic Reviews 2011(1):Mr000031. doi: 10.1002/14651858.MR000031.pub2 [published online first: 2011/01/21]

[17] Dwan K, Gamble C, Williamson PR, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLOS ONE 2013;8(7):e66844. doi: 10.1371/journal.pone.0066844 [published online first: 2013/07/19]

[18] Jones CW, Misemer BS, Platts-Mills TF, et al. Primary outcome switching among drug trials with and without principal investigator financial ties to industry: a cross-sectional study. BMJ Open 2018;8(2):e019831. doi: 10.1136/bmjopen-2017-019831

[19] Shepshelovich D, Yahav D, Tibau A, et al. Assessment of frequency and reporting of design changes among clinical drug trials published in influential medical journals.

European Journal of Internal Medicine 2020;71:45-49. doi: https://doi.org/10.1016/j.ejim.2019.08.009

[20] Boutron I, Dutton S, Ravaud P, et al. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. Jama 2010;303(20):2058-64. doi: 10.1001/jama.2010.651 [published online first: 2010/05/27]

[21] Chiu K, Grundy Q, Bero L. 'Spin' in published biomedical literature: A methodological systematic review. PLOS Biology 2017;15(9):e2002173. doi: 10.1371/journal.pbio.2002173

[22] Jellison S, Roberts W, Bowers A, et al. Evaluation of spin in abstracts of papers in psychiatry and psychology journals. BMJ Evidence-Based Medicine 2019 doi: 10.1136/bmjebm-2019-111176 [published online first: 2019/08/07]

[23] de Vries YA, Roest AM, de Jonge P, et al. The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: the case of depression. Psychological Medicine2018;48(15):2453-55. doi: 10.1017/s0033291718001873 [published online first: 2018/08/03]

[24] Leykin Y, DeRubeis RJ. Allegiance in psychotherapy outcome research: Separating association from bias. Clinical Psychology: Science and Practice 2009;16(1):54-65. doi: 10.1111/j.1468-2850.2009.01143.x

[25] Munder T, Brutsch O, Leonhart R, et al. Researcher allegiance in psychotherapy outcome research: an overview of reviews. Clinical Psychology Review2013;33(4):501-11. doi: 10.1016/j.cpr.2013.02.002 [published online first: 2013/03/19]

[26] Cuijpers P, Karyotaki, E., Ciharova, M., Quero, S., Pineda, B., Munoz, R., Figueroa, C. A meta-analytic database of randomised trials on psychotherapies for depression. January 27, 2020 doi: https://doi.org/10.17605/OSF.IO/825C6

[27] Vos T, Allen C, Arora M, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet 2016;388(10053):1545-1602. doi: 10.1016/S0140-6736(16)31678-6

[28] Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS medicine 2006;3(11):e442. doi: 10.1371/journal.pmed.0030442

[29] Cuijpers P. Four decades of outcome research on psychotherapies for adult depression: An overview of a series of meta-analyses. Canadian Psychology/Psychologie canadienne 2017;58(1):7-19. doi: 10.1037/cap0000096

[30] Cuijpers P, van Straten A, Smit F, et al. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. American Journal of Psychiatry 2008;165(10):1272-80. doi: 10.1176/appi.ajp.2008.07091422 [published online first: 2008/09/04]

[31] Cuijpers P. Meta-analyses in mental health research: A practical guide. 2016. [Available from: https://indd.adobe.com/view/5fc8f9a0-bf1e-49d3-bf5f-a40bfe5409e0 accessed March 2020]

[32] Gewandter JS, McKeown A, McDermott MP, et al. Data interpretation in analgesic clinical trials with statistically nonsignificant primary analyses: An ACTTION systematic review. The Journal of Pain 2015;16(1):3-10. doi: https://doi.org/10.1016/j.jpain.2014.10.003

[33] Boutron I, Ravaud P. Misrepresentation and distortion of research in biomedical literature. Proceedings of the National Academy of Sciences 2018;115(11):2613-19. doi: 10.1073/pnas.1710755115

[34] Cristea IA, Naudet F. Increase value and reduce waste in research on psychological therapies. Behaviour Research and Therapy 2019;123:103479. doi: https://doi.org/10.1016/j.brat.2019.103479

[35] Munder T, Fluckiger C, Gerger H, et al. Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? a meta-analysis. Journal of Counseling Psychology 2012;59(4):631-7. doi: 10.1037/a0029571 [published online first: 2012/09/06]

[36] Lieb K, von der Osten-Sacken J, Stoffers-Winterling J, et al. Conflicts of interest and spin in reviews of psychological therapies: a systematic review. BMJ Open 2016;6(4):e010606. doi: 10.1136/bmjopen-2015-010606 [published online first: 2016/04/28]

[37] Nosek BA, Lakens D. Registered reports: A method to increase the credibility of published results. Social Psychology 2014;45(3):137-41. doi: 10.1027/1864-9335/a000192

[38] Nutu D, Gentili C, Naudet F, et al. Open science practices in clinical psychology journals: An audit study. Journal of Abnormal Psychology 2019;128(6):510-16. doi: 10.1037/abn0000414

[39] Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332. doi: 10.1136/bmj.c332

[40] Center for Open Science: Simple registered report protocol preregistration 2020 [Available from: http://osf.io/rr/ accessed March 2020]

[41] An innovative way to publish. Nature 2019;571(7766):447. doi: 10.1038/d41586-019-02277-1 [published online first: 2019/07/25]

[42] Hardwicke T, Ioannidis J. Mapping the universe of registered reports. Nature Human Behaviour. 2018; 2 (11): 793–6

**Tables**

Table 1

*Characteristics of registrations and published trials: registration time point and characteristics of registered primary outcomes; trial registration status; and characteristics of trial and reported primary outcomes*

| **characteristics** | | **n (%)** |
|---|---|---|
| A) registration report of 134 registered trials | | |
| time point of registration | - prospective | 78/134 (58%) |
| | - retrospective | 56/134 (42%) |
| registered POs | - at least one PO mentioned | 134/134 (100%) |
| | - no PO mentioned | 0/134 (0%) |
| | - number of registered PO | 197 |
| quality of registered PO definition | - exactly defined | 51/197 (26%) |
| | - inexactly defined | 146/197 (74%) |
| B) trial report of 196 published randomized trials | | |
| registration status | - registered | 134/196 (68%) |
| | - not registered | 62/196 (32%) |
| trial effectiveness | - effective | 67/196 (34%) |
| | - not effective | 77/196 (39%) |
| | - not rateable[1] | 52/196 (27%) |
| reported POs | - at least one PO mentioned | 150/196 (77%) |
| | - no PO mentioned | 46/196 (23%) |
| | - number of reported POs | 194 |
| quality of reported PO definition | - exactly defined | 59/194 (30%) |
| | - inexactly defined | 135/194 (70%) |
| significance of PO | - significant | 89/194 (46%) |
| | - non-significant | 94/194 (48%) |
| | - not rateable[1] | 11/194 (6%) |

*Note.* PO = Primary Outcome.

[1] in some cases, we were not able to identify a primary outcome or a similar equivalent to assess statistical significance.

Table 2

*Forms of spin in 72 trials showing at least one non-significant primary outcome*

| **Spin** | **No. (%)** | |
|---|---|---|
| Spin in results section | **Abstract** (n = 70) | **Main Text** (n = 72) |
| Some type of spin | 51 (73%) | 44 (61%) |
| • Selective Reporting | 4 (6%) | 0 (0%) |
| • Distracting with secondary analyses | 15 (21%) | 4 (6%) |
| • Distracting with within-group differences | 11 (16%) | 4 (6%) |
| • Focus on secondary analyses | | |
| - Secondary analyses mentioned first | 15 (21%) | 19 (26%) |
| - Effect estimates mentioned for secondary analyses only | 12 (17%) | NA |
| - Only secondary analyses are presented in figures | NA | 7 (10%) |
| • Focus on within-group differences | | |
| - Within-group differences mentioned first | 15 (21%) | 29 (40%) |
| - Effect size mentioned for within-group differences only | 9 (13%) | NA |
| - Only within-group differences are presented in figures | NA | 1 (1%) |
| Spin in discussion section | | **Main Text** (n = 72) |
| Some type of spin | NA | 58 (81%) |
| • Selective reporting | NA | 4 (6%) |
| • Distracting with secondary analyses | NA | 15 (21%) |
| • Distracting with within-group differences | NA | 14 (19%) |
| • Focus on secondary analyses | | |
| - Secondary analyses mentioned first | NA | 22 (31%) |
| • Focus on within-group differences | | |
| - Within-group differences mentioned first | NA | 24 (33%) |
| • Interpreting non-significant primary results in a superiority trial as showing treatment equivalence | NA | 10 (14%) |
| • Claiming or emphasizing beneficial effect of treatment | NA | 28 (39%) |
| Spin in conclusion section | **Abstract** (n = 69) | **Main Text** (n = 48) |
| Some type of spin | 56 (81%) | 36 (75%) |
| • Selective reporting | 37 (54%) | 15 (31%) |

| | | |
|---|---|---|
| • Distracting with secondary outcomes | 8 (12%) | 11 (23%) |
| • Distracting with within-group differences | 7 (10%) | 7 (15%) |
| • Focus on secondary analyses | | |
| - Secondary analyses mentioned before | 4 (6%) | 5 (10%) |
| • Focus on within-group differences | | |
| - Within-group differences mentioned before | 3 (4%) | 1 (2%) |
| • Interpreting non-significant primary results in a superiority trial as showing treatment equivalence | 4 (6%) | 3 (6%) |
| • Claiming or emphasizing beneficial effect of treatment | 30 (43%) | 22 (46%) |

**Appendix (web-only)**

eTable 1: *Characteristics of the 196 included studies.*

eTable 2: *Registered and reported primary outcomes and rating of protocol discrepancies in the 196 included trials.*

eTable 3a: *Occurrence of spin in abstract sections for every non-significant primary outcome.*

eTable 3b: *Occurrence of spin in main text sections for every non-significant primary outcome.*

**Highlights**

- protocol discrepancies and spin in psychotherapy outcome research have not been investigated in detail so far
- protocol discrepancies are less frequent in psychotherapy trials which are registered prospectively as compared to retrospectively registered trials
- registration of psychotherapy trials is not associated with less spin in the publications

**Declarations of interest**

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf. MS, AM and LH declared that they had received salary from the Volkswagen Foundation. KL declared that he received a research grant by Volkswagen Foundation. However, any influence of the funder on the study is excluded. JB, PC, ND and JK have nothing to disclose. All authors declared: no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced submitted work.